Application of PCA Analysis and QR Decomposition to Address RFM's Ill-Posedness

Amin Alizadeh Naeini, Sayyed Hamed Alizadeh Moghaddam, Mohammad Moein Sheikholeslami, and AliReza Amiri-Simkooei

Abstract

Rational function model (RFM) is the most widely used sensor model in the remote sensing community. However, it suffers from ill-posedness, challenging its feasibility. This problem, i.e., ill-posedness, is mainly caused due to highly correlated coefficients of the RFM, which magnifies any small perturbations of observations, such as noise and instrumental error. This paper outlines a novel two-step method, called principal component analysis (PCA)-RFM, based on the integration of PCA and QR decomposition. In the first step, the PCA-RFM reduces the observational perturbations from the design matrix using the PCA. In the next step, the RFM's coefficients are estimated using a QR decomposition with column pivoting and least square method. According to the results, the PCA-RFM is less sensitive than its rivals to the changes of the ground control point (GCPs) distribution. Geometrically speaking, in addition, PCA-RFM is more accurate than recently established methods even in the presence of the small number of GCPs.

Introduction

Rational function model (RFM) is the most famous mathematical model that precisely defines the mathematical relation between the two-dimensional image and three-dimensional world coordinate systems. This model is convenient due to its simplicity, generality, and ease of implementation (Cao *et al.* 2017). However, the RFM generally faces ill-posedness and overparametrization problems mainly due to the existence of unnecessary and highly correlated RFM's parameters, known as rational polynomial coefficients (RPCs) (Moghaddam *et al.* 2018b).

Addressing the overparameterization and ill-posedness problems of the RFM has gained lots of attention in the literature. The existing methods in this context can be broadly categorized as regularization-based techniques and variable selection methods. ℓ_2-norm regularization technique, called Tikhonov or Ridge estimation (RE) method, is one of the first methods applied by Tao and Hu (2001) to overcome the problem of ill-posedness in the RFM. ℓ_1-norm regularization technique is another regularization-based technique, which was presented by Long, Jiao, and He (2015) in the context of the RFM. This technique, which is based on convex optimization algorithms, results in a vector of RPCs with some nonzero elements.

In addition to the abovementioned regularization-based techniques, some variable selection methods have been introduced in the literature. Although all these methods pursue an identical objective to select an optimal subset of RPCs, they apply different methodologies. In this regard, we can mention the methods based on nested regression (Tengfei, Weili, and Guojin 2014), stepwise-then-orthogonal regression (Li *et al.* 2017), statistical solutions (Moghaddam *et al.* 2017; Moghaddam *et al.* 2018b), and metaheuristic optimization algorithms (Gholinejad, Naeini, and Amiri-Simkooei 2018; Li *et al.* 2018; Moghaddam, Mokhtarzade, and Moghaddam 2018a; Naeini *et al.* 2017).

In addition to the aforementioned methods, some methods have been proposed that do not belong to the variable selection and regularization-based categories. In this regard, Zhou, Jiao, and Long (2012) replaced the least-squares (LS) estimation method with a Levenberg-Marquardt algorithm. Recently, Cao and Fu (2018) put forward a method based on a truncated singular value decomposition and the LS method. In that method, the design matrix is firstly decomposed into one diagonal matrix and two orthogonal matrices. Then the LS method is applied to estimate the unknown RPCs.

The problem of ill-posedness is quite severe when experimental data is inexact (e.g., noisy), which is always the case in practice because observational errors are inherent in the data. In the case of ill-conditioning, any small perturbation of data results in a significant error in the solution. In other words, the ill-posedness problem can magnify minor errors in the observations (Aytas, Afacan, and Tuna 2017).

Keeping this in mind, we propose a novel two-step method, based on the integration of principal component analysis (PCA) and QR decomposition. The proposed method mitigates mainly the ill-conditioning problem of the RFM by reducing the effects of data perturbation on the design matrix. To this end, in the first step, PCA initially transforms the design matrix to the principal component (PC) space. Secondly, in the PC space, the unnecessary components that have a considerably small variance are excluded. This makes the design matrix noise-reduced because the excluded components contain almost noise and have no essential information. Then, the noise-reduced design matrix is transformed back to its original space utilizing an inverse PCA transformation. Finally, this design matrix is used to estimate the unknown RPCs.

Because the exclusion of the unnecessary components causes the noise-reduced design matrix to be rank deficient, a novel estimation method is applied in the second step of the method. This estimation method, which is based on a LS method and the QR decomposition with column pivoting, handles not only the rank deficiency but also the problem of overparametrization.

Because the effects of the noise are reduced in the design matrix, the estimated RPCs are less affected by the ill-conditioning problem. From now on, we call the proposed method PCA-RFM.

Amin Alizadeh Naeini is with the Dept. of Earth and Space Science and Engineering, York University, 4700 Keele St, Toronto, ON M3J1P3, Canada (naeini@yorku.ca)

Mohammad Moein Sheikholeslami and AliReza Amiri-Simkooei are with the Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, Isfahan 8174673441, Iran.

Sayyed Hamed Alizadeh Moghaddam is with the Faculty of Geodesy and Geomatics Engineering, Khajeh Nasir Toosi University of Technology, Tehran 1996715433, Iran.

Photogrammetric Engineering & Remote Sensing
Vol. 86, No. 1, January 2020, pp. 17–21.
0099-1112/20/17–21
© 2020 American Society for Photogrammetry and Remote Sensing
doi: 10.14358/PERS.86.1.17

Methodology

The RFM is expressed as follows (Ma 2013; Tao and Hu 2001):

$$l = \frac{P_1(E,N,h)}{P_2(E,N,h)}, \quad s = \frac{P_3(E,N,h)}{P_4(E,N,h)} \tag{1}$$

where P_i is a third-order polynomial, (l,s) are the image coordinates of a point and (E, N, h) are its coordinates in the object space. Note that the zero-order term in P_2 and P_4 are considered to be "1", and the coordinates of the points are commonly normalized in the range of [–1, 1] in both image and object spaces. Equation 1 is commonly reformulated as follows (Long, Jiao, and He 2015):

$$P_1(E, N, h) - lP_2(E, N, h) = 0 \tag{2}$$

$$P_3(E, N, h) - sP_4(E, N, h) = 0 \tag{3}$$

The unknown RPCs ($\boldsymbol{x}$), can be estimated using the ordinary least-squares (OLS) regression as $\boldsymbol{x} = (A^TA)\,A^T\boldsymbol{y}$, where A is the design matrix and $\boldsymbol{y}$ is the vector of observations. However, the OLS solution is affected by the problems of ill-posedness and overparameterization.

Addressing the previously mentioned problems, we put forward the PCA-RFM method. The PCA-RFM generally consists of two steps. In the first step, which itself consists of two phases, the design matrix (A) is initially transformed into a new space, namely, PCs. Then, in the final phase, some inessential PCs of A, containing almost noise, are excluded, and a new noise-reduced design matrix is reconstructed. In the last step, this reconstructed design matrix is used to solve the unknown RPCs. To do this, we apply the QR decomposition with column pivoting and the LS method (Golub and Van Loan 2012) because the reconstructed design matrix is rank deficient. The steps of the proposed method are delineated below.

Step 1: Construct the design matrix $A_{m\times n}$, applying GCPs. $m = 2k$ is the total number of observation equations, where k is the number of applied GCPs, and $n = 78$ is the total number of unknown RPCs. See Long, Jiao, and He (2015), and Moghaddam *et al.* (2018b) for more information of how the design matrix is calculated.

Step 2: Transform the matrix A to the PC space as follows:

Step 2-1: Compute the mean-centered $\bar{A}$ to facilitate the further calculation (see Equation 7).

$$\bar{A} = A_i - \text{mean}\,(A_i) \tag{7}$$

where A_i is the ith column of the design matrix A, i.e., $A = [A_i]$.

Step 2-2: Calculate the covariance matrix (C) of the new design matrix $\bar{A}$ by applying Equation 8:

$$C = \frac{\bar{A}^T\bar{A}}{m-1} \tag{8}$$

Step 2-3: Conduct an eigen value decomposition on C:

$$C = V\Omega V^T \tag{9}$$

$$V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]_{n\times n} \tag{10}$$

$$\Omega = \text{diag}([\lambda_1, \lambda_2, \ldots, \lambda_n]) \tag{11}$$

where V is a matrix in which the columns are eigen vectors ($\boldsymbol{v}_i$) of C, and Ω is a diagonal matrix that contains the corresponding eigen values (λ_i), i.e., $\Omega_{ii} = \lambda_i$. Note that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$.

Step 2-4: Transform the mean-centered $\bar{A}$ to the PC space to get S as follows:

$$S = \bar{A}V \tag{12}$$

Step 3: Partition the matrices S and V in Equation 12 as follows:

$$S = [S_1 \vdots S_2] = \bar{A}[V_1 \vdots V_2] = [\bar{A}V_1 \vdots \bar{A}V_2] \tag{13}$$

where the $78\times r$ matrix V_1 contains r eigen vectors that their corresponding eigen values are larger than a threshold t. Moreover, the matrix V_2 contains the remaining eigen vectors having eigen values smaller than the threshold.

The columns of S_2 are the PCs with considerable small variances (i.e., eigen value). These PCs contain almost unnecessary information (such as noise) and can exert a remarkable influence on the derived products in the presence of the ill-posedness problem. Therefore, PCA-RFM weeds out the S_2 and only considers $S_1 = \bar{A}V_1$. In this paper, we empirically set $t = 0.01$.

Step 4: Back-transform S_1 to its original space using an inverse PCA transformation, and reconstruct the noise-reduced design matrix $\bar{A}_{\text{nr}}$ as follows:

$$\bar{A}_{\text{nr}} = S_1V_1^T = \bar{A}\ V_1V_1^T \tag{14}$$

The LS estimation step demands the real noise-reduced design matrix. To get that, the columns of $\bar{A}_{nr}$ are required to be decentered because $\bar{A}_{nr}$ is a mean-centered matrix (see Step 2-1). This is conducted by adding the corresponding mean values, which are calculated in Step 2-1.

$$A_{\text{nr}_i} = \bar{A}_{\text{nr}_i} + mean(A_i) \tag{15}$$

The matrix A_{nr} has a rank of r. This indicates that some of the columns of the new design matrix A_{nr} are strictly dependent, resulting in a rank deficiency. Therefore, we have faced with a rank deficient LS problem if we use A_{nr} as the design matrix.

Step 5: Estimate the unknown RPCs utilizing the "basic solution" of the QR decomposition with column pivoting and the LS method, described in (Golub and Van Loan 2012). Finally, the PCA-RFM results in a vector of RPCs with some nonzero components, which are expected to be less affected by the problems of ill-posedness and overparameterization.

Delivered by Ingenta
IP: 127.0.0.1 On: Wed, 13 November 2019 07:24:41
Copyright: American Society for Photogrammetry and Remote Sensing

Experiments

Data Sets

In these experiments, nine real data sets (HJ, L5, PL, GE, WV, S1A, S1B, IK, and IRS) from eight different platforms (Table 1) were used to evaluate the performance of the proposed method.

According to Table 1, images with various ground sampling distances (GSDs), ranging from 0.5 m for PL, GE, and WV to 30 m for L5 and HJ were used in the experiments. Regarding the geometric level of correction, all the images except for S1A experienced a primary level of correction, including the correction of distortions introduced by Earth rotation and curvature, resampling of the images to a uniform GSD, and assigning the images to a map projection system. The S1A data set is a geometrically raw image with only a radiometric correction phase.

The control points (CPs) of five data sets, i.e., PL, GE, WV, IK, and IRS, the accuracy of which is better than one meter, were chosen from distinct features of 1:2000 digital reference maps. For the two Spot data sets, a differential global positioning system (DGPS) technique was applied, and finally, for HJ and L5, an automatically matching technique was employed to

extract the CPs from a geo-referenced Landsat-5 image (Long, Jiao, and He 2015).

Competing Methods

Since the proposed method has some prominent qualities, and due to its remarkable performance, we chose methods which can compete with the proposed method in different aspects, including 1) the capability in dealing with the ill-posedness and overparametrization problems, and 2) the applicability in the presence of a small number of GCPs (e.g., ten GCPs). Accordingly, the competing methods, ranging from classical to the cutting-edge, were selected, and their abbreviations to increase the convenience and concision, are brought as follows:

1. OLS: ordinary least squares (Tao and Hu 2001)
2. RE: ridge estimation (Tao and Hu 2001)
3. LM: Levenberg-Marquardt (Zhou, Jiao, and Long 2012)
4. TSVD: truncated singular value decomposition (Cao and Fu 2018)
5. USS-RFM: uncorrelated and statistically significant RFM (Moghaddam *et al.* 2018b)
6. L1LS: ℓ_1-norm regularized technique with least squares estimation (Long, Jiao, and He 2015)

Regarding the parameter setting of our proposed method, we empirically set the threshold t to 10^{-2}, providing reasonable results for all data sets. The RE and L1LS methods has a regularization parameter that must be set in prior. According to (Long, Jiao, and He 2015, Tao and Hu 2001), we set these parameters to $10^{-5}/k$ (k is the total number of GCPs) and $10^{-2.5}$ for the L1LS and RE, respectively. The parameters of the USS-RFM method were set according to (Moghaddam *et al.* 2018b) as $\gamma = 10^{-6}$ and $\alpha = 0.2$. To implement the L1LS using Lasso via the LARS algorithm, we used the linear_Mmodel module from the freely available Scikit-Learn Python package (Pedregosa *et al.* 2011).

Discussion and Results

In this section, we give and then discuss the results of different analyses in three separate subsections. The results of the normal case analysis (NCA) are given in the first subsection. The GCP distribution analysis (GDA) results are reported in the second subsection. The results of the limited number of GCP analysis (LGA) are provided in the last subsection.

Normal Case Analysis

In this analysis, 40 GCPs (which is the least number of GCPs for having a positive degree of freedom) of each data set were selected randomly and the rest served as independent checkpoints (ICPs). Therefore, only data sets with the minimum required GCPs (i.e., 40) could be used. It means that S1A and S1B were not used here. The results of this analysis are more reliable than two other analyses due to its positive degree of freedom. Root-mean-square-error (RMSE) was calculated over the ICPs as a quality assessment measure. Table 2 shows the results of the NCA.

From Table 2, the first point is the results of the OLS, which are not satisfying at all. This unsatisfactory performance of the OLS, in line with (Long, Jiao, and He 2015, Moghaddam *et al.* 2018b), can be justified by the problems of ill-posedness and overparameterization, which challenge the application of the RFM. Note that in the OLS, no actions are taken to address those problems. This also shows that it is necessary to mitigate these pernicious problems before using RFM.

According to Table 2, OLS has the worst performance in terms of accuracy. Meanwhile, the USS-RFM, L1LS, and the PCA-RFM have an excellent performance in all data sets. The arithmetic mean of RMSEs provided by the proposed method is 0.88 pixel that is lower than those of the L1LS (1.09 pixel) and the USS-RFM (1.12 pixel). In order to statistically prove the superiority of the PCA-RFM over the L1LS and USS-RFM, we applied a Mann-Whitney U test.

Table 1. Data set specifications, including GSD, number of CPs, and the reference sources of these points.

Data Set	Platform	GSD (m)	Reference data	Area Type	No. of CPs
HJ	HJ-1	30	Landsat-5	Plain	200
L5	Landsat-5	30	Landsat-5	Mountainous, Hilly	200
PL	Pleiades	0.5	Maps	Urban	70
GE	GeoEye-1	0.5	Maps	Urban	70
WV	WorldView-3	0.5	Maps	Urban	65
S1A	Spot-3 1A	10	DGPS	Rural, Mountainous	28
S1B	Spot-3 1B	10	DGPS	Rural, Mountainous	28
IK	IKONOS	1	Maps	Urban	74
IRS	IRS-P5	2.5	Maps	Semiurban	77

Table 2. NCA report of the proposed PCA-RFM and the competing methods, including OLS, RE, LM, TSVD, USS-RFM, and L1LS. The best result in each data set is given in bold.

Data Set	Methods (RMSE/Pixel)						
	OLS	RE	LM	TSVD	USS-RFM	L1LS	PCA-RFM
HJ	46671.2	4.88	7.16	9.37	0.84	**0.63**	0.64
L5	57588.53	1.27	7.51	1.27	1.56	0.59	**0.59**
PL	8075.27	1.66	24.65	1.65	**0.85**	1.13	0.91
GE	1154.6	1.03	130.01	1.29	0.93	**0.66**	0.95
WV	1071.31	3.76	17.89	2.98	**0.72**	1.64	0.72
IK	19995.93	0.92	4.22	0.98	1.67	**0.85**	1.04
IRS	21492.45	5.3	13.16	2.47	**1.28**	2.14	1.33
Average	22292.76	2.69	29.23	2.86	1.12	1.09	**0.88**

The Mann-Whitney U test is a statistical procedure used to compare two sets of samples. Each of the PCA-RFM, L1LS, and USS-RFM methods was run ten times to generate a statistically sufficient population. Subsequently, the Mann-Whitney U test was used in order to determine whether the RMSE values resulted from PCA-RFM were significantly different from those of the USS-RFM and L1LS methods.

The test proved that the improvement, resulted from the PCA-RFM, is statistically significant over L1LS and USS-RFM even at the 99% confidence interval. This statistical proof, along with the fact that the average RMSE of the PCA-RFM over the data sets is less than those of the USS-RFM and L1LS, verify the superior performance of the proposed method.

A possible reason for the superiority of the proposed method is the proposed noise reduction phase. As previously mentioned, the ill-posedness problem magnifies the small perturbations of data, caused by random noise or instrumental error, for example. PCs with small eigen values contain noise and undesirable data perturbations. In the proposed method, such PCs are excluded from the design matrix, which reduces the effect of the ill-posedness problem. This is in contrast to the other methods where no action is taken to reduce the data perturbations that are magnified without addressing the ill-posedness problem. Additionally, the applied LS method based on the QR decomposition results in a vector of RPCs with few nonzero elements. Thus, the problem of overparametrization is addressed as the proposed method excludes some unnecessary RPCs from the model.

Regarding the computational time, all the methods were equally fast and were run in a small fraction of a second.

GCPs Distribution Analysis

In the GDA, a modified fivefold cross-validation (MFFCV) was applied to comprehensively evaluate the efficiency of the methods in various distributions of GCPs and ICPs. The MFFCV,

Delivered by Ingenta
IP: 127.0.0.1 On: Wed, 01 Jan 2020 00:00:00
Copyright: American Society for Photogrammetry and Remote Sensing

firstly, splits the CPs of each data set into five distinct subsets with approximately the same dimension. Then, in each fold, it allocates one of the five subsets as GCPs and the rest as the ICPs. In each one of the folds, RMSE on ICPs is calculated, and eventually, the average RMSE and standard deviation (STD) over the five folds are reported (Table 3).

In contrast to the regular fivefold cross-validation, in the MFFCV, a higher number of ICPs are used in each fold. Hence, we can better analyze and judge the sensitivity of the methods to the GCP distribution changes. Note that similar to the NCA, S1A, and S1B were not used in the GDA. Due to the poor performance of RE, LM, and TSVD methods in the NCA, these methods were not included in the GDA.

As is apparent from Table 3, the PCA-RFM except for three data sets (i.e., HJ, PL, and IK) has the best performance, which indicates its ability to overcome various distributions significantly. For HJ and IK, the L1LS method has a little bit better performance than the PCA-RFM. However, on average of seven data sets, the PCA-RFM leads to respectively 25.89% and 40.01% improvement in terms of the RMSE and the STD, compared to the L1LS. Considering the results of the USS-RFM, the proposed PCA-RFM achieved a decrease in the RMSE values by 17.90%, on average. It can be concluded that, compared to the competing methods, the proposed method has impressive resistance to changes in the distribution of the GCPs and it is less sensitive than its competitors are.

Note that the result of the USS-RFM in GE case study is marked by an asterisk (*). This means that the USS-RFM could not identify an acceptable RFM in some folds of the MFFCV. In this case, the unacceptable folds were discarded. This also shows that the USS-RFM method is more sensitive to the changes in the GCP distribution than the L1LS and PCA-RFM methods.

Table 3. GDA report of the proposed PCA-RFM, USS-RFM, and L1LS. The best result in each data set is given in bold, and the result of the USS-RFM in the GE case study is marked by an asterisk.

	Methods (mean RMSE in Pixels ± STD)		
Data Set	**USS-RFM**	**L1LS**	**PCA-RFM**
HJ	0.86 ± 0.05	**0.65 ± 0.09**	0.69 ± 0.09
L5	1.39 ± 0.47	0.67 ± 0.08	**0.61 ± 0.03**
PL	**1.00 ± 0.15**	2.31 ± 1.39	1.17 ± 0.22
GE	1.18 ± 0.21*	1.70 ± 0.61	**0.98 ± 0.09**
WV	1.14 ± 0.37	4.29 ± 2.81	**1.05 ± 0.08**
IK	1.86 ± 0.15	**0.95 ± 0.08**	1.39 ± 0.19
IRS	2.02 ± 1.20	3.95 ± 1.64	**1.69 ± 0.18**
Average	1.35 ± 0.37	2.07 ± 0.95	**1.08 ± 0.12**

Table 4. LGA report of the proposed PCA-RFM, USS-RFM, and L1LS. The best result in each data set is given in bold, and the result of the USS-RFM in the GE case study is marked by an asterisk.

	Methods (mean RMSE in Pixel ± STD)		
Data Set	**USS-RFM**	**L1LS**	**PCA-RFM**
HJ	1.14 ± 0.34*	**0.95 ± 0.05**	1.26 ± 0.12
L5	3.70 ± 1.22*	1.49 ± 0.37	**1.33 ± 0.16**
PL	1.48 ± 0.73	1.46± 0.73	**1.15 ± 0.23**
GE	1.07 ± 0.15*	1.45 ± 0.33	**1.15 ± 0.34**
WV	1.71 ± 0.46	3.70 ± 2.47	**1.23 ± 0.19**
IK	5.11 ± 3.98	**1.22 ± 0.22**	1.49 ± 0.19
IRS	2.91 ± 3.28	8.13 ± 12.95	**1.78 ± 0.30**
S1A	1.18 ± 0.37*	3.83 ± 3.29	**1.14 ± 0.28**
S1B	0.91 ± 0.01*	2.24 ± 1.23	**1.14 ± 0.26**
Average	2.13 ± 1.17	2.71 ± 2.40	**1.30 ± 0.23**

doi:10.14358/PERS.86.1.001 Delivered by Ingenta
Copyright American Society for Photogrammetry and Remote Sensing
On: Wed, 03 Nov 2021 10:24:31

Limited Number of GCPs Analysis

The objective of the LGA is to evaluate the performance of the proposed and competing methods when a small number of GCPs are available. To accomplish this, we randomly selected ten GCPs. Furthermore, this selection was made five times to challenge the performance of the methods for changes in the GCP distribution. Results of this analysis are given in Table 4, where mean and STD of RMSEs are reported in mean ± STD format. Similar to the GDA, RE, LM, and TSVD methods did not take part in the LGA.

From Table 4, on average, the PCA-RFM has shown 29.29% and 39.57% improvement in the average RMSEs and STDs, respectively, compared to the L1LS. Considering the second and fourth columns of Table 4, the PCA-RFM method leads to 20.52% improvement, on average, in terms of the mean RMSE values compared to the USS-RFM. Similar to the previous section, the results of the USS-RFM in some case studies are marked by an asterisk. Because acceptable results could not be achieved by this method in some rounds of the LGA.

In the LGA, the competing methods provided a poor performance in comparison to the PCA-RFM. It supports a conclusion that the PCA-RFM achieves more accurate results than the rivals when few GCPs are available. In addition, the dependency of the proposed method on the GCP distribution is less than the other methods as it resulted in smaller STD values.

Conclusion

In this paper, we have presented a novel two-step method based on the integration of PCA and QR decomposition, mainly attempting to deal with the ill-posedness problem. In the first step, the PCA-RFM method uses the PCA to reduce the noise and other data perturbations from the design matrix, leading to a noise-reduced but rank deficient design matrix. In the second step, the PCA-RFM applies the QR decomposition to solve the rank deficient LS problem and estimate the unknown RPCs.

Evaluating the feasibility of the PCA-RFM, we designed three analyses: NCA, GDA, and LGA. Geometrically speaking, the NCA showed that PCA-RFM was about 7% and 16% more geometrically accurate than the L1LS and USS-RFM methods, respectively. Note that this improvement was statistically significant at a 99% significance level. In the GDA, the method's dependency on GCP distribution was investigated, using a fivefold cross-validation approach. The results proved that the proposed method was less sensitive to changes in the GCP distribution.

Finally, the performance of the PCA-RFM and the other competing methods in the presence of small numbers of GCPs was evaluated in the LGA. LGA's results demonstrated that the PCA-RFM was geometrically 29.29% and 20.52% more accurate than the L1LS and USS-RFM methods, respectively. From the LGA, it can be concluded that the PCA-RFM is a more effective solution in the case that a small number of GCPs are available.

As with any new approach, there are some unresolved issues that may present challenges over time. One might be the setting of the threshold t in the first step of the PCA-RFM. The automatic selection of t could be a topic for future research.

Acknowledgements

The authors would like to thank Dr. T. Long from the Chinese Academy of Sciences, Beijing, China, for his help in the implementation of L1LS method and providing HJ and L5 data sets. They would like to thank Prof. M. J. V. Zoej from the K. N. Toosi University of technology, Tehran, Iran, and Dr. S. Homayouni from the University of Ottawa, Ottawa, ON, Canada for providing some data sets in the case studies.

References

Aytas, N., E. Afacan and E. Tuna. 2017. Analysis of the incoming signal on the antenna array using singular value decomposition. Pages 18–21 in *Proceedings of ISER 56th International Conference*, held in Rome, Italy.

Cao, J., J. Fu, X. Yuan and J. Gong. 2017. Nonlinear bias compensation of ZiYuan-3 satellite imagery with cubic splines. *ISPRS Journal of Photogrammetry and Remote Sensing* 133:174–185.

Cao, J. and J. Fu. 2018. Estimation of rational polynomial coefficients based on singular value decomposition. *Journal of Applied Remote Sensing* 12 (4):044003.

Gholinejad, S., A. A. Naeini and A. Amiri-Simkooei. 2018. Robust particle swarm optimization of RFMs for high-resolution satellite images based on K-fold cross-validation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Golub, G. H. and C. F. Van Loan. 2012. Matrix Computations. JHU Press.

Li, C., X. Liu, Y. Zhang and Z. Zhang. 2017. A stepwise-then-orthogonal regression (STOR) with quality control for optimizing the RFM of high-resolution satellite imagery. *Photogrammetric Engineering & Remote Sensing* 83 (9):611–620.

Li, C., J. Liu, X. Wang, X. Liu and Y. Wu. 2018. Stepwise-then-intelligent algorithm (STIA) for optimizing remotely sensed image rectification. *International Journal of Remote Sensing*: 1–20.

Long, T., W. Jiao and G. He. 2015. RPC estimation via l_1 -norm-regularized least squares (L1LS). *IEEE Transactions on Geoscience and Remote Sensing* 53 (8):4554–4567.

Ma, R. 2013. Rational function model in processing historical aerial photographs. *Photogrammetric Engineering & Remote Sensing* 79 (4):337–345.

Moghaddam, S. H. A., M. Mokhtarzade, A. A. Naeini and Moghaddam, S. A. A., 2017. Statistical method to overcome overfitting issue in rational function models. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 42:23–26.

Moghaddam, S. H. A., M. Mokhtarzade and S. A. A. Moghaddam. 2018a. Optimization of RFM's structure based on PSO algorithm and figure condition analysis. *IEEE Geoscience and Remote Sensing Letters* 15 (8):1179–1183.

Moghaddam, S. H. A., M. Mokhtarzade, A. A. Naeini and A. Amiri-Simkooei. 2018b. A statistical variable selection solution for RFM Ill-posedness and overparameterization problems. *IEEE Transactions on Geoscience and Remote Sensing* 56 (7):3990–4001.

Naeini, A. A., S. H. A. Moghaddam, S. M. J. Mirzadeh, S. Homayouni and S. B. Fatemi. 2017. Multiobjective genetic optimization of terrain-independent RFMs for VHSR satellite images. *IEEE Geoscience and Remote Sensing Letters* 14 (8):1368–1372.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12 (Oct):2825–2830.

Tao, C. V. and Y. Hu. 2001. A comprehensive study of the rational function model for photogrammetric processing. *Photogrammetric Engineering and Remote Sensing* 67 (12):1347–1358.

Tengfei, L., J. Weili and H. Guojin. 2014. Nested regression based optimal selection (NRBOS) of rational polynomial coefficients. *Photogrammetric Engineering & Remote Sensing* 80 (3):261–269.

Zhou, Q., W. Jiao and T. Long. Solution to the rational function model based on the Levenberg-Marquardt algorithm. Pages 2795–2799 in *Proceedings of the 2012 9th IEEE International Conference on Fuzzy Systems and Knowledge Discovery (FSKD).*

Delivered by Ingenta
IP: 127.0.0.1 On: Wed, 03 Nov 2021 10:29:31
Copyright: American Society for Photogrammetry and Remote Sensing